



Data File Catalog

Dmitry Litvintsev, Fermilab

Eric Wicklund, Fermilab

CDF Data Handling Review, September 25, 2001

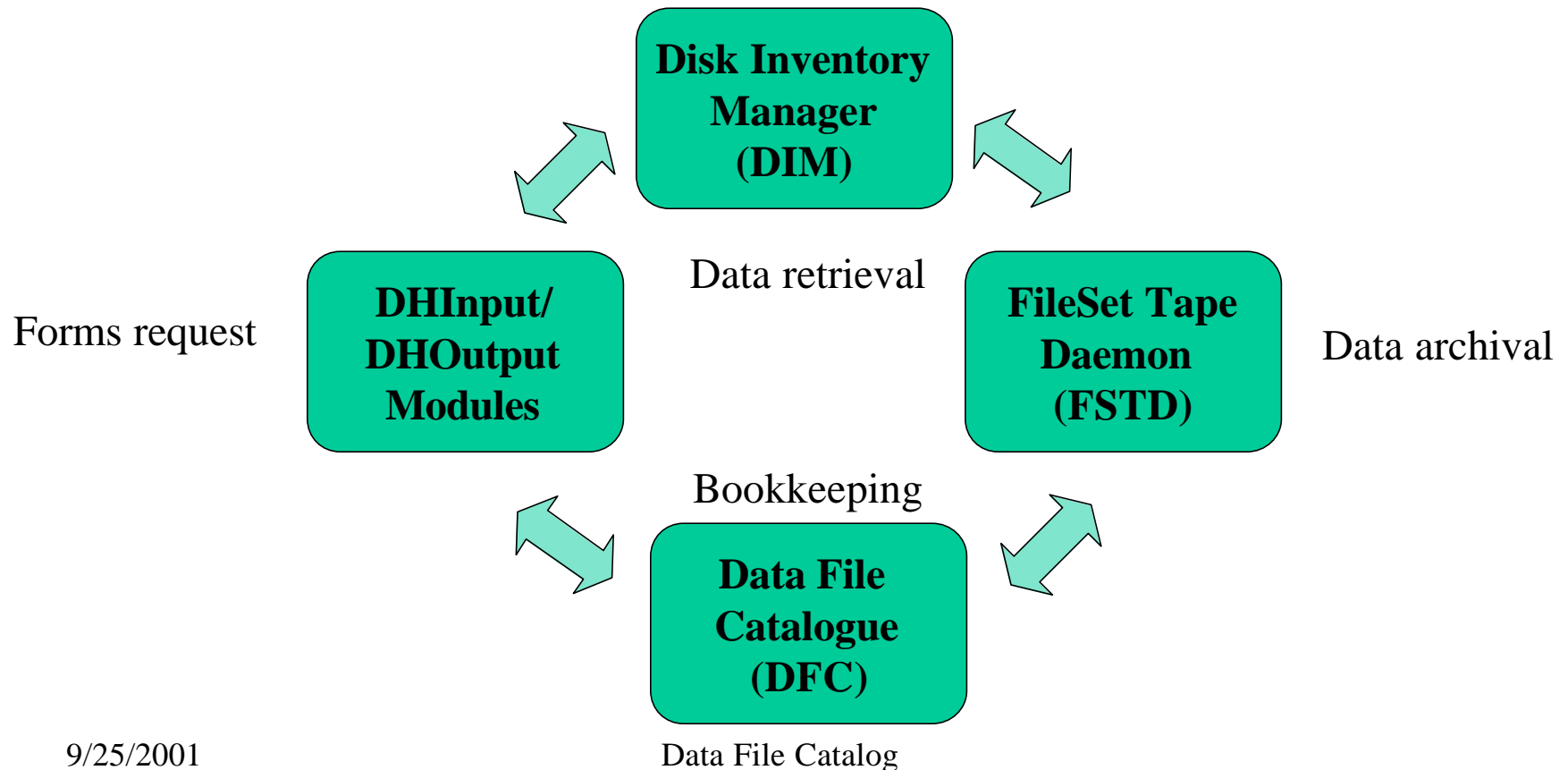
- Introduction
- Component Status
- Secondary Datasets
- Conclusion

DH System (simplified)

- DH System:

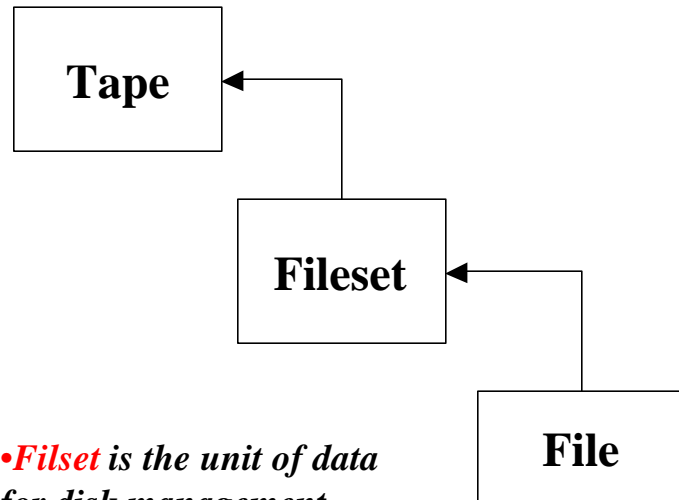
Organize, access and archive the data

- Components of DH System:



Hierarchical Data Organization

Physical (DH) view



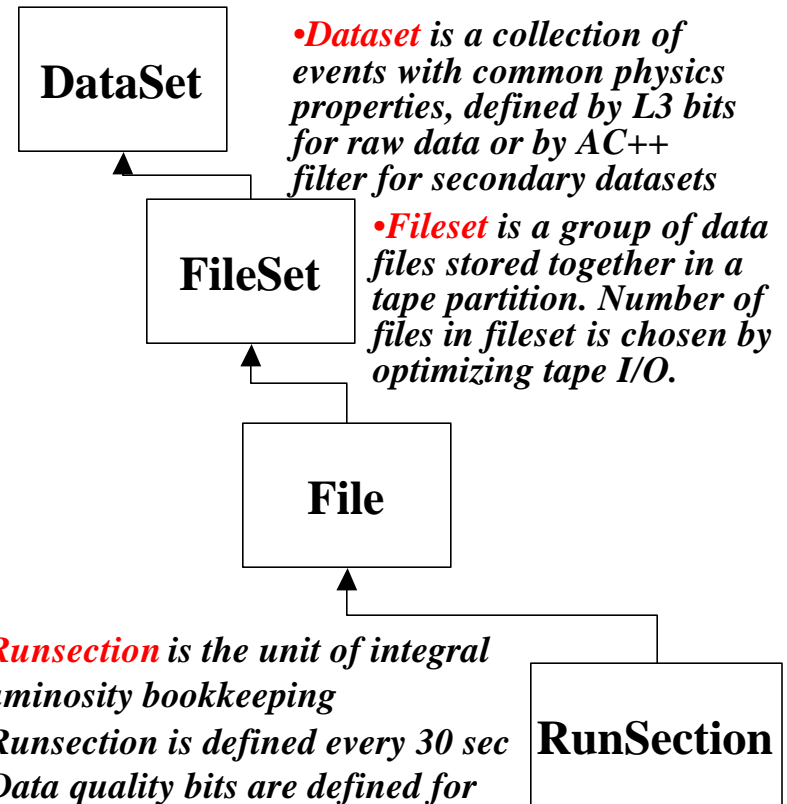
•**Filset** is the unit of data for disk management

•Output files of derived streams must be aligned to runsection boundaries (all events of any runsection should be in one file)

•Datasets are grouped into streams in such a way that minimizes the number of events that are stored in more than one stream

•A stream can be either primary (Raw data) or derived (reconstructed data, PAD etc.)

User view



•**Dataset** is a collection of events with common physics properties, defined by L3 bits for raw data or by AC++ filter for secondary datasets

•**Fileset** is a group of data files stored together in a tape partition. Number of files in fileset is chosen by optimizing tape I/O.

•**Runsection** is the unit of integral luminosity bookkeeping

•Runsection is defined every 30 sec

•Data quality bits are defined for Runsection

•Dynamic trigger prescales change on Runsections' boundaries

What is Data File Catalog

**Data File Catalog (DFC) = Relational Database (DB) +
+ Application Program
Interface (API)**

Data File Catalog

Information, necessary to locate any CDF Dataset (primary and derived), is stored in relational database. It contains:

- **Four core tables** corresponding to four elements of data access hierarchy:

CDF2_DATASETS	- datasets
CDF2_FILESETS	- filesets
CDF2_FILES	- files
CDF2_RUNSECTIONS	- runsections

- It also contains general bookkeeping information like **tape** volume contents and allocation:

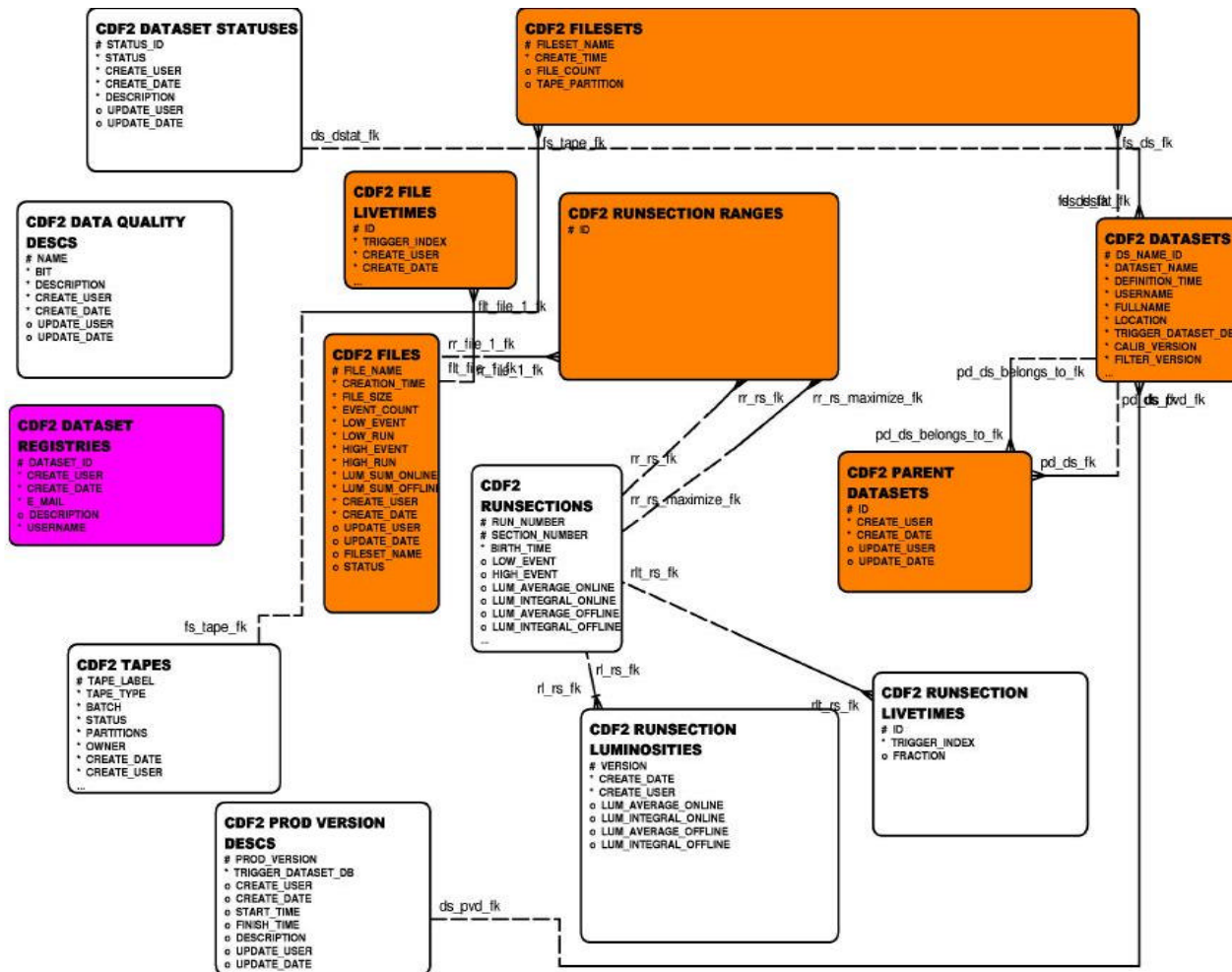
CDF2_TAPES

- Keeps track of **parent datasets**:

CDF2_PARENT_DATASETS

- physics related bookkeeping information such as data quality, trigger and filter used, average and integral luminosities etc
- Data access granularity of DFC is runsection. I.e. individual events are not described.
- Central DFC at Fermilab is an Oracle DB. Remote sites may use Msql to set up their local DFC.
- **DataFileDB** package provides all the code necessary to manipulate DFC from a C++ program or from command line. The package currently supports interfaces to Oracle and Msql implementations of DFC.

DFC DB schema



9/25/2001

Data File Catalog

Some Numbers

Table Name	Initial Mb	Next Mb	min	max	#rows	row size B	Used Mb
CDF2_DATASETS	1.1	1.1	1	20	84	195	0.016
CDF2_FILESETS	15.2	15.2	1	5	5567	51	0.273
CDF2_FILES	667	667	1	5	111189 (56197)	112	11.91
CDF2_RUNSECTION_RANGES	494	494	1	5	115309	50.4	5.547
CDF2_RUNSECTIONS	242	242	1	5	526905	54.4	27.34
CDF2_TAPES	4.7	4.7	1	5	1136	72.1	0.078

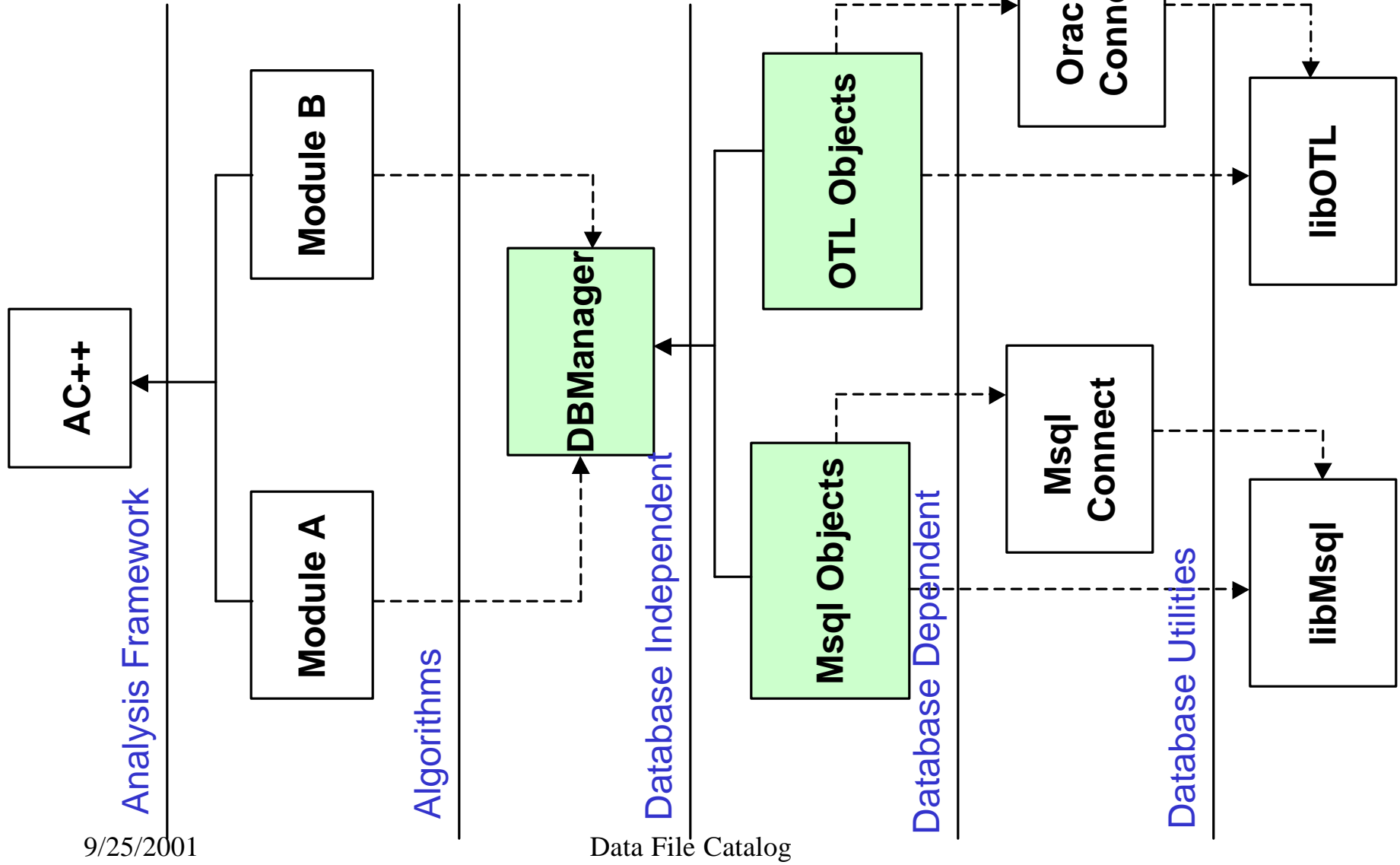
Oracle DB

- Oracle DB schema has been developed using Oracle Designer tool.
- It has passed review by Fermilab DBAs (J.Trumbo, N.Stanfield, A.Kumar)
- Space assignment for DFC tables and indices is sufficient to keep 4 years worth of data
- Space allocation is monitored on daily basis:

http://wwwcdserver.fnal.gov/cd_public/ods/db_stats/data/db_stats.html

Schema is stable. No major changes are foreseen

C++ API



DataFileDB

- DataFileDB provides DB interface layer for managing the mapping of persistent data stored in DB tables to transient C++ objects seen my event processing software
- Allows keyed access to data in the DB
- Places get/put/change methods on top of transient class for retrieval or storage of objects of this class using a key

DataFileDB

- DataFileDB has five basic classes that define rows of DB tables:
 - DFCDataset
 - DFCFileset
 - DFCFile
 - DFCRunsection
 - DFCTape
- These objects have necessary interfaces and utility classes allowing user to manipulate them

DataFileDB:Example

Example of retrieval of all files from dataset aphysr from DB identified by “production”

```
typedef Manager<DFCFiles,DFCFileKey> DFCFiles_mgr;  
DFCFiles_mgr m("production","DFCFiles");  
//key class associated with file  
DFCFileKey key;  
key.setDatasetNameID("aphysr");  
//typedef std::vector<DFCFile> DFCFiles  
//typedef Handle<DFCFiles> DFCFiles_var  
DFCFiles_var files;  
m.get(key,files)
```

DataFileDB Status

DFC C++ API provides interfaces for all relevant tables. Development has been completed. It is now in maintenance mode

Secondary Datasets proposal

User Books Development

- An idea to use Oracle-based DFC by physics group and individuals for bookkeeping secondary datasets has been put forth in CDF 5380
- The following requirements have been identified:
 - The physics groups and individual users should have the same 24x7 service provided to official users for writing datasets
 - Any user in CDF must be able to read datasets created by another user
 - The physics groups must not be permitted to delete entries made by official users while being able to delete entries made by themselves
 - The physics groups and individual users must not be permitted to fill up Oracle tablespace of the official users
 - Physics groups must not delete entries made by other physics groups
 - Physics groups must not delete entries made by individuals
 - An individual user must not be able to delete entries made by other users

Secondary Datasets proposal

- It is suggested that there will be some tables common for the whole experiment, such as CDF2_TAPES table which holds list of tapes in a tape robot and some other tables for which physics groups or users may have personalized versions.

- Six such tables have been identified:

CDFPEXO.CDF2_FILES

CDFPEXO.CDF2_FILESETS

CDFPEXO.CDF2_DATASETS

CDFPEXO.CDF2_PARENT_DATASETS

CDFPEXO.CDF2_FILE_LIVETIMES

CDFPEXO.CDF2_RUNSECTION_RANGES

} exotic book

- Make use of Oracle quota and access mechanisms

Secondary Datasets procedures

Procedures

- Application for **externally authenticated** Oracle account
- This account will have necessary privileges (**grants**) in order to be able to setup private version of DFC
- User will then need to execute a simple script:

DFCMakeUserBook joseph [/@cdfopfprd](#)

This script will create a private DFC schema in JOSEPH account

- User may use DHOutput/DHInput modules to fill/read his private DFC. For example:

```
talk DHInput
  include dataset "My Jpsi sample" book=joseph
exit
```


Secondary Datasets registry

Dataset Registry GUI

The screenshot shows a Netscape browser window titled "Netscape: DFCTools" displaying the "Dataset Registry Tool" interface. The browser's address bar shows the URL: <http://ncdf68.fnal.gov/cgi-bin/dfctools/index.php?username=litvinse&op=DatasetRegistryTool&crypt=zq8zAfR7mNe>. The interface includes a menu bar (File, Edit, View, Go, Communicator) and a toolbar with buttons for Back, Forward, Reload, Home, Search, Netscape, Print, Security, Shop, and Stop. Below the toolbar, there are links for "Dataset Registry Tool" and "What's Related". The main content area is titled "Dataset Registry Tool [Main Menu | Create | Locate]" and contains a "Create" section with the following fields: "Book(Group) Name" (litvinse (8637)), "Stream" (z), "Production Pass" (z), "Database" (Development), and "Description" (empty). A "Submit" button is located below the "Description" field. At the bottom of the browser window, there is a footer that reads "Written by Benjamin Tsai and Pegen Shen and [Dmitry Litvinse](#)" and "Fermi National Accelerator Laboratory".

datasetID is generated automatically from unique userID and supplied stream and production version characters

9/25/2001

Data File Catalog

Secondary Datasets booking

The screenshot shows a Netscape browser window titled "Netscape: DFCTools". The address bar displays the URL: <http://ncdf68.fnal.gov/cgi-bin/dfctools/index.php?username=litvinse&op=DatasetTool&crypt=GzE1AD10lhDsY6G&r>. The browser's menu bar includes File, Edit, View, Go, Communicator, and Help. The toolbar contains icons for Back, Forward, Reload, Home, Search, Netscape, Print, Security, Shop, and Stop. The bookmarks bar shows several links, including "www.sesameworkshop.org/", "Internet", "Lookup", "New&Cool", "www.cnn.com/STYLE", "The ROOT System Home Page", and "cdfcodebrowser.fnal.gov/CdfCode".

The main content area is titled "Dataset Tool" with links for "Main Menu", "Create", and "Locate". Under the "Create" section, there is a form with the following fields:

- Database:
- Book(Group) Name:
- Parent ID:
- Dataset ID:
- Description:
- Production Version:
- Dataset Name:
- Filter Version:
- Trigger Version:
- Calibration Version:
- Location:

A "Submit" button is located below the form fields.

At the bottom of the page, there is a footer that reads: "Written by Benjamin Tsai and Pegen Shen and Dmitry Litvinsev" with a link to "Local Notices". To the right of the footer, it says "Fermi National Accelerator Laboratory".

9/25/2001

Data File Catalog

Secondary Dataset Status

DFC Part

- All scripts are ready
- The DataFileDB package and DHMods package have been modified to accept **book** parameter
- Dataset Registry procedure has been worked out and tested
- Easy to use WWW GUI interface exists for Datasets
- **Need to implement similar interface for Tape allocation**

Operational Issues

- DFC maintenance (relatively minor load)
- There is a problem of file entries (about 55,000) which were not recorded on tape. There needs to be a mechanism of clean deletion of those entries
- At some point revisit CSL-LumMon interaction when filling on-line luminosity numbers
- It would be good to have an AIT-2 tape drive attached to fcdfora1 for regular backups of DFC
- Operational issues related to user books:
 - setting up and testing communication between File Set Tape Daemon and DFC when allocating and writing a tape.
 - Error handling within FSTD s/w
 - user consulting
 - user support
 - Oracle operations
 - scripts to copy/move datasets from one book to another
 - installation and support of DFC at the remote institutions

Conclusion

- CDF DFC has been functioning smoothly for about a year
- We have not experienced any significant problems with both Oracle and C++ software
- DFC Software and Oracle schema are ready to handle secondary datasets